心理科学进展 2023, Vol. 31, No. 4, 608–621
Advances in Psychological Science

https://doi.org/10.3724/SP.J.1042.2023.00608

言语想象的神经机制*

王勇丽[1] 葛胜男[2] Lancy Lantin Huang[3] 万 勤[1] 卢海丹[1]

([1]华东师范大学教育学部康复科学系，上海 200062) ([2]华东师范大学生命科学学院，上海 200241)
([3]爽琅国际贸易有限公司，上海 200131)

摘 要 言语想象不仅在大脑预处理机制方面起到重要的作用，还是目前脑机接口领域研究的热点。与正常言语产生过程相比，言语想象的理论模型、激活脑区、神经传导路径等均与其有较多相似之处。而言语障碍群体的言语想象、想象有意义的词语和句子时的脑神经机制与正常言语产生存在差异。鉴于人类言语系统的复杂性，言语想象的神经机制研究还面临一系列挑战，未来研究可在言语想象质量评价工具及神经解码范式、脑控制回路、激活通路、言语障碍群体的言语想象机制、词语和句子想象的脑神经信号等方面进一步探索，为有效提高脑机接口的识别率提供依据，为言语障碍群体的沟通提供便利。

关键词 言语想象，发音运动计划，脑机接口，神经机制
分类号 B842

1 引言

言语产生是指人们将思想通过语音形式表达出来的过程，包括概念化、言语组织和发声三个阶段(张清芳，杨玉芳，2003；张清芳 等，2021)，其中前两个阶段属于言语产生的计划过程，而发声阶段则属于言语执行过程(蔡笑，张清芳，2020)。言语执行是一个涉及发音运动计划、发音运动执行和感觉反馈的复杂过程(Guenther, 1995, 2006)，该过程需要涉及中枢神经系统和外周神经系统，若神经系统及其传导路径损伤，则会出现言语执行过程受阻，诸如脑卒中、帕金森病、脑外伤、脑瘫、小脑萎缩等脑损伤患者常会伴随言语执行功能障碍。

言语想象(speech imagery)是一种对自己或他人言语事件产生的内部类知觉(quasi-perceptual)体验(Orpella et al., 2022)，是不发出任何声音或做出面部动作的内在发音的一种状态(Torres-García et al., 2016)。言语想象是“我们内心的小声音”，是自上而下的认知加工过程，在科学和哲学等领域都有悠久的历史(MacKay, 1992; Solms, 2000; Weber & Bach, 1969)。由于言语想象不会受到肌肉运动的干扰，便于脑功能检测技术的信号采集，被认为是一个潜在的研究大脑预处理机制的宝贵工具(Keller & Mrsic-Flogel, 2018)。

视觉想象(Pearson, 2019)、听觉想象(Gu et al., 2019)、触觉想象(Yoo et al., 2003)、嗅觉想象(Djordjevic et al., 2005)和运动想象(白学军 等, 2016; Souto et al., 2020)等感官或运动方面的研究表明，想象与实际感官执行的神经表征有高度重叠。但是言语想象既包含了对发音器官的运动想象(想象说)，又包含了对语音感知的想象(想象听)，是有别于单纯的感知想象和运动想象的一种特殊的双重机制的想象方式(Tian et al., 2016)。近年来，神经心理学家和语言学家探索了言语想象的神经机制，有些学者认为言语想象可以激活与实际言语产生中发音运动计划相似的脑区(Cooney et al., 2018; Naito et al., 2002; Tian & Poeppel, 2012)，与言语执行中的自我监控(Tian & Poeppel, 2015)、感觉预测和反馈机制(Hickok et al., 2011; Jones & Fernyhough, 2007)有相同之处，有些学者甚至认为言语想象可能从高层次的语言表征开始就与言语产生有相似之处(Oppenheim & Dell, 2010)。另外，言语想象神经传导路径方面的研究检测到了

收稿日期：2022-06-14
* 国家社科基金青年项目(19CYY019)。
通信作者：王勇丽, E-mail: wylkangfu@126.com

微弱的面部肌肉表面肌电信号(Ma et al., 2019; Orpella et al., 2022), 认为言语想象与言语执行的神经传导路径也相似。基于言语想象与言语执行神经机制的相似性, 言语想象疗法被认为是脑损伤导致的言语障碍群体康复治疗的新技术(窦智等, 2021; Durand et al., 2018; Laures- Gore et al., 2021)。而有些研究则认为言语想象与外显言语产生(overt speech)具有不同的神经信号特征, 不能简单的将言语想象认为是言语执行中的一部分过程(Proix et al., 2021; Stark et al., 2017)。厘清言语想象的神经机制, 有助于我们对言语产生机制有更清楚的认识(Keller & Mrsic-Flogel, 2018; Orpella et al., 2022), 也能够为言语障碍群体提供更加高效的康复治疗技术。因此, 本文的第一个主要内容是梳理言语想象与言语产生神经机制的异同。

不同言语想象内容的神经机制是当前言语想象需要突破的技术难题, 该部分的研究主要集中在脑机接口领域(brain-computer interface, BCI) (Cooney et al., 2018)。BCI 是通过对脑活动信号的采集与处理来实现人或动物的大脑控制外接设备的一种通讯系统, 为言语障碍和肢体障碍群体与外界沟通带来了极大的便利(陈霏, 潘昌杰, 2020; Cooney et al., 2018)。言语想象应用于 BCI 的识别率和舒适性优于运动想象和视觉想象(Nguyen et al., 2017), 是一种新型的交际型脑机接口方式。BCI 领域对言语想象的研究, 强调对复杂言语想象内容脑神经信号解码的准确性。提高神经信号解码率的前提是了解言语想象时具有最佳解码能力的大脑激活区域(刘艳鹏　等, 2022)和采用提取到最大信息量的技术或方法(Cooney et al., 2018; Proix et al., 2021)。皮层脑电(electrocorticographic, ECoG)、脑电图(electroencephalography, EEG) (Cooney, et al., 2018)、脑磁图(magnetoencephalography, MEG) (Orpella et al., 2022)和功能磁共振(functional magnetic resonance imaging, fMRI)等技术促进了言语想象状态下神经信号的采集和解码。已有文献综述指出言语想象的内容也可能是影响其神经信号的主要因素之一。常见的想象内容有元音想象(vowel imagery)、音节想象(syllable imagery)、单词想象(word imagery)和句子想象(sentence imagey)等(陈霏, 潘昌杰, 2020; Cooney et al., 2018), 内容的分类特征和其附加的语言信息会产生不同的神经信号。本文的第二个主要内容是梳理不同想象内容对应的脑神经信号特征。

2　言语想象与言语产生神经机制的异同

言语产生是一个极其复杂的过程, 言语想象的神经机制研究多是以正常言语产生的神经机制作为对照(Cooney et al., 2018)。学者们分别从言语产生的理论模型、脑区激活和神经传导路径、言语障碍群体等方面对比了言语想象的神经机制。

2.1　言语想象与言语产生有相似的的理论模型

当前流行的言语产生模型有 DIVA (Directions Into Velocities of Articulators)模型(Guenther, 1994, 1995, 2006)和状态反馈控制(state feedback control, SFC)模型(Heremans et al., 2011; Hickok, 2012; Houde & Nagarajan, 2011)。DIVA 模型阐述了人体言语产生的过程, 包括前馈系统的运动指令(自上而下)、反馈系统的听觉和体感反馈错误纠正指令(自下而上), 以及联系前馈和反馈系统的内部前馈模型(internal forward models)三个环节(Guenther, 1994, 1995, 2006)。其中, 内部前馈模型是将运动指令副本传递中枢神经系统预测即将接受到的体感和听感刺激的信息传递至反馈系统(蔡笑, 张清芳, 2020; Guenther & Vladusich, 2012)。大量研究表明, 言语想象与实际言语产生过程有部分的重叠, 强调了与发音器官的运动计划和体感、听感结果预测的过程高度重叠。一项 MEG 和 fMRI 的研究认为, 言语想象的过程类似于 DIVA 模型中言语产生的前馈和反馈系统均暂停工作, 只有内部前馈模型发挥重要作用的过程(Tian & Poeppel, 2010), 即预测了体感和语音听感的结果(Kilteni et al., 2018)。而这个预测的过程是按照“运动计划(motor planning)→运动指令副本(motor efference copy)→第一前馈模型运动估计(顶叶皮层)→听觉指令副本(perceptual efference copy)→第二前馈模型听觉估计(感觉皮层)”的顺序串行(Tian & Poeppel, 2010; Tian et al., 2016)。运动假说(motor hypothesis)认为, 言语想象是实际言语产生过程的衰减版, 具有明确的发音运动计划(articulatory plan) (Cooney et al., 2018)。行为学的研究也表明, 言语想象与言语产生的过程高度相似, 是运动系统对感觉结果的预测(Scott et al., 2013)。SFC 模型强调了言语产生过程中存在一个内部言语计划反馈回路和一个外部错误监测反馈回路(Hickok et al., 2011; Hickok,

2012; Houde & Nagarajan, 2011)。Orpella 等人(2022)采用MEG技术对言语想象状态下的脑功能活动进行解码，发现在言语想象的开始阶段检测到运动前区和颞顶区的功能活动，这与 SFC 理论的内部言语计划反馈回路结果一致，接着在想象500ms 以后听觉和双侧运动区的活动明显激活，与 SFC 外部错误监测反馈回路一致(Orpella et al., 2022)。值得注意的是，以上研究强调了言语想象与言语执行的相似之处，均未提及言语想象是否包含正常言语产生的概念化、言语组织等高层次认知加工过程。

抽象假说(abstraction hypothesis)认为言语想象不仅会涉及到实际言语产生中的发音运动计划，还包括高层次语言表征的加工过程，其与实际言语产生中言语器官动作执行之前的所有过程都相同(Oppenheim & Dell, 2010)。例如，研究表明言语想象涉及了与重要认知功能相关的联想语言学习(Orpella et al., 2021)、预测过程和语音工作记忆(Alderson-Day & Fernyhough, 2015; Perrone-Bertolotti et al., 2014)等方面，说明言语想象涉及了更高层次的认知加工过程。

以上研究均表明了言语想象与实际言语产生过程相似或部分重叠，不同之处在于言语想象并未激活最后一步发音运动执行的环节。这些理论支持了将言语想象应用于 BCI 的可能性，也是当前替代运动想象和视觉想象而成为 BCI 的主要手段的依据(Nguyen et al., 2017)。然而，目前较少有理论来阐述二者之间的差异。在言语障碍群体的应用和 BCI 的研究中，学者依据行为学和脑电信号的差异提出二者可能存在不同的加工机制。例如，失语症患者的实际言语产出障碍明显，但其仍旧保留较好的言语想象的能力(Fama et al., 2017; Stark et al., 2017)。一项对比言语想象和言语产生状态脑电信号识别率的研究表明，在实际言语产生时高频宽带(broadband high-frequency, BHA)信号识别率最高，而在言语想象时低 θ 波、低 β 波和低 γ 波的识别率最高，甚至超过了言语执行，这说明言语执行和言语想象在某些加工机制方面有一定的差异(Proix et al., 2021)。甚至有研究表明，言语想象比实际言语产出更容易受到左岛盖部的影响(Geva et al., 2011)。

以上争议可能与言语想象的研究还面临很多挑战有关。在研究方法上的不足，体现在由于言语想象没有明确的行为表现，通过行为学的比较研究较为局限(Alderson-Day & Fernyhough, 2015; Martin et al., 2018)，后续可以探索言语想象质量评价的行为学方法。在神经信号解码上的不足，有些脑损伤患者大脑语言中枢皮层网络受损，只能对残存的其他区域的神经信号进行解码(Guenther et al., 2009; Wilson et al., 2020)。某种情况下这些区域可能存在一些代偿功能，或者解剖位置使得神经信号的解码很难实现，导致与正常群体的神经信号有差异。

2.2 言语想象与言语产生激活的脑区相似

言语想象激活的脑区与实际言语产生有很多相似之处，主要体现在发音运动计划和听觉中枢的区域，但言语想象激活的脑区信号强度较弱。

言语想象的 fMRI、EEG 和 MEG 等研究表明，言语想象激活了与发音运动计划相似的脑区，包括布洛卡区(Broca's area)、辅助运动区(supplementary motor ares, SMA)、运动前区(premotor cortex)、脑岛(insula)和小脑(cerebellum)等区域。早期，研究者为了降低发音时面部肌肉运动造成的 fMRI 伪影，尝试采用事件相关的功能磁共振(event-related fMRI)方法来降低言语状态下脑激活信号的伪影(Birn et al., 1999; Palmer et al., 2001)，结果表明受试者想象发单词与实际大声读出单词的大脑激活信号有相似的地方(Palmer et al., 2001)。后续有大量 fMRI 的研究表明言语想象激活了左额下回后部(布洛卡区) (André et al., 2005; Hurlburt et al., 2016; Shergill et al., 2001)，辅助运动区、运动前区、岛叶(André et al., 2005; Cooney et al., 2018; Shergill et al., 2001)等。MEG 相关的研究也表明言语想象会激活以上区域(Tian & Poeppel, 2010, 2012)，这些区域是正常言语产生时主要负责发音运动计划的区域(Bohland & Guenther, 2006; Guenther et al., 2006; Guenther & Vladusich, 2012; Kearney & Guenther, 2019)。另外，研究表明言语想象和实际言语产生时均激活了小脑(Ackermann & Hertrich, 2003; Naito et al., 2002)，且小脑在发音运动计划中也扮演了重要的角色(Kearney & Guenther, 2019; Parrell et al., 2017)。这些结论可以用理论模型中提及的言语想象有明确的发音运动计划来解释。“第一第二前馈模型”、Orpella 等人的研究，均强调了言语想象过程中涉及对构音器官运动的体感预测结果，且运动假说

和抽象假说也强调了言语想象涉及了言语产生中的发音运动计划阶段。

正常言语产生过程中颞叶皮层主要负责将听觉反馈和运动信息进行整合(Hickok, 2012), 言语想象同时包括基于体感−运动的“想象说”和基于记忆提取的“想象听”, 其过程同样激活了颞叶皮层。Riaz 等人(2015)探索了不同数据分析方法下元音实际发音和言语想象状态的 EEG, 发现二者在位于韦尼克区的信号相似性较高。Courson 和 Tremblay (2020)的荟萃分析表明尤其是颞中回后部参与了言语想象。言语想象的近红外光谱成像(functional near infrared spectroscopy, fNIRS)研究也表明, 通过言语想象来操控脑机接口时左侧颞叶和颞顶叶皮层被激活(Sereshkeh et al., 2018)。Tian 等人的研究采用先进的 MEG 和 fMRI 技术, 且在实验设计上更加精细地区分了言语构音动作想象(articulation imagery, AI)和听觉想象(hearing imagery, HI)状态脑区激活的异同, 发现在额叶−顶叶感觉运动系统中, AI 时更多激活了与发音运动计划相关的脑区, 激活顶叶的 150~170ms 后, 还观察到双侧颞叶的活动(Tian & Poeppel, 2012); HI 时更多激活了与听觉记忆恢复有关的额中叶、顶叶下皮层和顶叶内沟等区域(Tian et al., 2016)。AI 和 HI 的双侧颞叶活动模式, 类似于实际听到声音时听觉相应的脑区激活(Tian & Poeppel, 2012)。因此, AI 和 HI 的神经表征既有分离又有一定范围的重叠, Tian 等学者将此称为言语想象的“双流预测模型” (Tian & Poeppel, 2012)。言语想象激活了听觉皮层与言语想象理论模型中“第一第二前馈模型”中提及的第二前馈模型听觉估计、Orpella 等人研究中提及想象 500ms 后听觉皮层激活一致。

言语想象可以激活与发音运动计划和听觉中枢的脑区, 而与实际言语产生相比激活信号强度较弱(Alderson-Day & Fernyhough, 2015; Pei et al., 2011; Proix et al., 2021)。Jahangiri 和 Sepulveda (2018)的研究表明言语想象时激活的脑电信号弱于言语产生, 且不足以解释言语信息的复杂性。另外, 言语想象并未引起肉眼可见的发音器官运动, 其是否激活了言语产生中运动执行的主要运动皮层(又称初级运动皮层 M1)还存在比较大的争议。有研究者认为, 言语想象激活了 M1, 但是激活信号比较弱, 因此不足以引起真正的运动(Ehrsson et al., 2003)。然而, 有些研究则认为想象过程中辅助运动区对 M1 区发挥了抑制作用, 并未检测到 M1 区的激活信号(高晴, 陈华富, 2010; Tak et al., 2015)。首先, M1 是否被激活, 可能与检测技术有关(Martin et al., 2018), 检测技术之间的敏感性存在一定的差异, 后续可以进一步进行检测技术的对照研究。其次, 关于 M1 区被抑制的证据, 也可能与受试者刻意控制运动有关。受试者被要求不引起实际运动, 言语想象实验过程中受试者内心会刻意控制运动执行, 因此 M1 区被抑制。再者, 言语想象同时包括了想象听和想象说, 当二者同时想象时, 可能会分散受试者的注意力, 从而影响了 M1 的激活, 后续可能需要严格的实验设计控制注意力的影响(Bruno et al., 2018)。

综上所述, 在脑区激活方面, 言语想象与实际言语产生有很多相似之处, 包括发音运动计划相关脑区和听觉皮层, 在进行言语想象相关的实验研究或 BCI 的识别时考虑应覆盖以上区域。与实际言语产生相比, 言语想象激活的信号强度较弱, 尤其在初级运动皮层的激活上争论较大, 后续可能需要进行不同检测技术之间的对照研究和严格的实验设计控制无关变量。

2.3　言语想象与部分言语产生的神经传导路径相似

言语产生过程的运动执行阶段, 需要将大脑皮层的信息借助相应的神经传导通路向下传递, 才能引起言语器官的协调运动(Bohland & Guenther, 2006; 蔡笑, 张清芳, 2020; Hickok, 2012)。学者们也探索了言语想象的神经传导路径, 多集中在采用 EEG、MEG 等检测大脑皮层之间的功能连接(function connectivity) (Sandhya et al., 2015)、采用肌电图(EMG)检测肌群的肌肉电信号(Lebon et al., 2008), 以及采用弥散张量成像技术(diffusion tensor imaging, DTI)追踪脑白质纤维束的通路(Li et al., 2018)。

MEG 的研究探索了言语想象过程中不同脑区之间的神经传导路径, 结果显示从左侧颞叶的活动, 到顶叶、前额岛叶、双侧运动前区的变化, 反映了言语想象时也存在与言语产生相似的听觉−运动皮层的功能连接和整合过程(Orpella et al., 2022)。Sandhya 等学者(2015)采用 EEG 检测了言语想象状态下脑区之间的神经传导路径, 表明额叶和颞叶大脑区域之间存在与言语产生相似的左

右半球之间的相互作用，同样是来自左半球的信息流比右半球大，且左侧颞叶跨电极相关性较高，但在言语执行时是左侧额叶跨电极相关性较高。采用表面肌电(sEMG)的研究则表明，在言语想象时可以精准地采集到面部肌肉的微弱电信号(Ma et al., 2019; Orpella et al., 2022)，这似乎说明言语想象存在与发音运动执行相似的神经传导路径。但有学者认为，sEMG 检测到的微运动是心理语言常见的现象，是运动执行的副产品而不能完全抑制(Perrone-Bertolotti et al., 2014)，是否可以用发音运动执行的神经通路解释还存在疑问。Orpella 等学者(2022)则认为，言语想象同样存在发音运动计划被执行的过程，只不过是到外周神经肌肉系统时被抑制了。因此，言语想象的神经传导路径研究还需要进一步探索。

以上研究表明，言语想象与部分言语产生的神经传导路径相似。然而，回顾言语产生相关的神经传导通路，包括锥体系、锥体外系、基底神经节控制回路、小脑控制回路等中枢神经系统传导，以及脑神经、脊神经等外周神经系统传导(Bohland & Guenther, 2006; Duffy, 2013)。以上研究仅涉及了两侧大脑皮层之间的传导路径，目前尚缺乏言语想象中锥体系、锥体外系、控制回路等相关的传导路径的证据。外周神经系统的研究也还存在争议，均需要进一步探索。目前磁共振的 DTI 技术追踪脑白质纤维已经在各类神经损伤的疾病中广泛应用，后续可以采用 DTI 技术进一步探索言语想象的神经传导路径。

2.4 言语障碍群体言语想象的神经损伤机制

言语想象是认知心理加工过程，以上所涉及的发音运动计划、听觉皮层和神经传导路径的神经解剖结构受损，可能会影响言语想象的能力，研究表明言语障碍群体的言语想象受损与言语产生的损伤程度并非完全正相关。

失语症通常是由于左侧颞叶—顶叶—额叶等网络受损导致的获得性语言障碍(Klingbeil et al., 2019; Schumacher et al., 2019)，这不仅影响局灶部位的功能，还会影响该部位与其他脑网络之间的联系(姚婧璠 等, 2021)。研究表明，失语症患者言语想象能力有一定程度的受损，尤其是想象同音词或押韵词方面存在困难，但很多其他内容的言语想象并不存在困难(Fama et al., 2019; Geva et al., 2011; Langland-Hassan et al., 2015)。例如，左侧大脑中动脉区域受损的失语症患者中，言语功能损伤较重的患者更依赖于言语想象(内心语言)来进行命名和书面图片的描述，表明言语想象还可能和言语产生中的工作记忆有相似的机制(Geva et al., 2011)，且有临床干预的研究通过言语想象来改善失语症患者的动词命名能力(Durand et al., 2018)和名词命名能力(Laures-Gore et al., 2021)，以及改善失语症的负面的心理状况(Barrows et al., 2021)、语言表达的流畅性(Barrows et al., 2021; Haire et al., 2021)和言语的流利性(窦智 等, 2021; 吴金香, 2019)等。Langland-Hassan 等人(2015)的研究则表明失语症(布洛卡失语、命名性失语为主)的患者名词命名能力要优于无声命名(言语想象)能力，且二者之间相关性不明显。而 Fama 等人(2019)的研究表明，失语症患者自我报告的言语想象与失语的严重程度相关，MRI 检测言语想象能力与大脑腹外侧尤其是感觉运动皮层的血流病变相关。以上失语症患者的言语想象与实际言语产出之间的差异，可能是与受损的神经解剖结构位置和严重程度不同有关(Fama et al., 2017)，目前尚缺乏研究报道不同类型和损伤程度失语症之间言语想象能力的差异和机制(Langland-Hassan et al., 2015)。

口吃患者的神经病理机制为脑 44 区与左顶叶之间的功能耦合缺陷(Neef et al., 2016)，调查研究表明，65%的口吃受访者表示在内心想象言语时没有口吃的情况(Netsell & Bakker, 2017)。临床应用研究表明想象哼唱旋律可以促进口吃患者的言语流利性(Neef et al., 2016; Neef et al., 2018)，而且想象流畅言语可以提高口吃患者的阅读广度(Arongna et al., 2020)。目前尚缺乏口吃患者言语想象和言语产生机制差异的研究，这些应用研究并不能很好的解释言语想象的机制，并且以上研究也多是将言语想象与其他疗法联合起来应用，缺乏更有价值的临床随机对照试验和循证医学证据证明言语想象疗法的治疗效果。

依据以上言语障碍群体的结论，结合上文所介绍言语想象时会涉及多个脑区和神经传导通路，相关神经的损伤可能使言语想象能力变慢、变弱，结合部分失语症患者言语想象能力并未受限等结论，说明局部性的神经损伤并不能使患者完全丧失言语想象能力。想象可以重塑脑功能(吴拾瑶 等, 2021)，言语想象治疗已经用于失语症和口吃

人群的言语障碍, 而言语想象涉及言语运动计划和听觉皮层, 以及神经传导通路, 这些区域受损会导致运动性言语障碍, 对于这类群体的言语想象机制和应用研究还较为缺乏, 值得进一步探索。

3　不同言语想象内容的脑神经信号

脑功能检测技术促进了人类快速认识言语产生的机制, 自 20 世纪 60 年代起学者们便逐渐探索了不同大脑皮层区域与言语任务之间的关系(张清芳, 杨玉芳, 2003), 不同语义范畴对应的脑区差异也比较大。借鉴言语产生的研究结论, 言语想象的研究也从探索想象不同元音时脑神经信号的差异, 逐渐发展到想象单词、辅音、句子及不同声调等脑神经信号的检测和解码(刘艳鹏 等, 2022)。

3.1　想象元音/音节时的脑神经信号

元音/音节是正常语言学习时较为简单的语音单位, 早期大量言语想象相关的研究倾向于采用元音或无意义音节作为想象内容, 其与正常言语产生有相似的脑神经信号。DaSalla 等人(2009)的研究指出, 在执行元音/a/和/u/想象的起始时, C3、CZ 和 C4 电极(国际 10~20 系统)位置出现负波趋势, 在 300 ms 左右出现正波, 并且这些波形与真实言语产生时的事件相关脑诱发电位(event-related potential, ERP)非常类似。杨晓芳和江铭虎(2014)的研究发现在执行 4 个元音和 4 个辅音想象任务时的 ERP 波形与真实发音器官运动引起的颅内及头皮电位时间进程相似。Orpella 等学者(2022)通过 MEG 探索了无意义音节层面(/pa/, /ta/, /ka/)言语想象的脑功能信号, 结果表明其与正常言语产生的差别不大。这些结论与前述言语想象的理论以及激活的言语计划和听觉相关脑区的结论一致, 且元音或无意义音节不涉及语义加工成分, 降低了大脑活动时脑电信号的复杂性(Ramirez-Quintana et al., 2021), 使其言语想象的脑电信号与正常言语产生相似。

3.2　想象汉语词语/英语单词时的脑神经信号

在进行赋有语义的汉语词语/英语单词的想象时, 与正常言语产生时的脑神经信号存在某种程度的差异。Proix 等人(2021)最大限度地增加语音表征、语义类别和数量的可变性, 包括对实词动词、抽象动词、实词名词和抽象名词等单词的语料进行言语想象, 结果表明正常言语产生时感觉运动皮层、颞叶皮层的 BHA 激活比较明显, 而言语想象时 θ 波、β 波、γ 波比较显著, BHA 不显著。Chengaiyan 等人(2020)探索了“辅音−元音−辅音”结构(consonant vowel consonant, CVC)单词的言语想象, 表明言语想象时颞叶的 θ 波占主导, 而在言语产生(发音)时, 额叶的 γ 波占主导。Nguyen 等人(2017)对比了想象状态下长单词、短单词和元音的脑信号来探究影响言语想象分类效果的因素, 包括发音时频特征和语料复杂程度等, 结果表明短单词和长单词在布洛卡区和韦尼克区出现时频差异, 短单词之间和音位之间的分类性能相似, 这表明影响言语想象分类效果的是发音时频特征; 短单词比长单词更容易在高频段(31~70Hz)和布洛卡区受到抑制, 但是长单词相比于短单词能提供更高的 Kappa 系数, 表明复杂程度越高的单词越容易利用脑神经信号进行区分。郭苗苗等人(2018)分别探索了想象四个不同词性和语义的汉字(喝, 右, 吃, 冷)时的脑电信号, 结果表明想象不同意义的汉字时在 α 波和 β 波随时间动态变化上有不同的表现, 反映了想象不同语义的汉字时大脑思维活动的差异。不同言语想象内容与脑神经信号的差异见表 1。这些差异一方面可能是由于语料的复杂程度和词性差异等涉及复杂的语言加工引起的, 另一方面也可能是研究者选择神经信号采集的脑区不同造成了差异。一些学者在研究言语想象的 BCI 信号采集时, 多考虑覆盖左侧额叶、颞叶和双侧感觉运动区(Jahangiri & Sepulveda, 2017; Pei et al., 2011; Proix et al., 2021; Sereshkeh et al., 2018; Wang et al., 2013)等区域, 也有多数研究者为了不遗漏较弱信号的区域而选择全脑区(Dash et al., 2020; 郭苗苗 等, 2018; Orpella et al., 2022)。虽然有研究表明仅利用左半脑通道信息的 BCI 系统的性能不差于利用全脑通道信息的 BCI 系统, 利用左半脑通道的信息足够可以提取出汉字发音想象的脑电特征(Wang et al., 2013), 但值得注意的是, 大部分研究均重点关注了激活信号较强的区域, 部分信号较弱的区域往往都不进行结果的报道, 信号较弱的脑区激活也可能是区分语义加工重要的区域。

从赋有语义的词汇产生时脑区激活机制可知, 言语想象的大脑加工机制非常复杂。BCI 的目的是将想象的言语尽可能正确的识别出来, 激活信号相对较弱的脑区可能也在 BCI 识别中扮演了重

要的角色。刺激方式会影响到脑神经电信号的采集，Zhang 等(2020)人探索了普通话四声调的言语想象，结果表明视听结合刺激的神经信号解码率(80.1%)比单纯视觉刺激时(67.7%)的高。后续可以对比信号特征提取的脑区和刺激方式方面进行进一步研究。

表 1　不同言语想象内容与脑神经信号的差异

研究	被试	想象对比	内容	覆盖脑区	工具	信号	结果
Proix et al. (2021)	13 例癫痫患者	言语想象 VS 言语产生(发音)	6 个单词：spoon, cowboy, battlefield, swimming, python, telephone	感觉运动皮层、颞上回、颞中回、颞下回、下额叶	ECoG (侵入性)	BHA、θ 波、β 波、γ 波	①言语想象时 BHA 不显著，θ 波、β 波、γ 波显著；②言语执行时以上脑区显著。
Cheng-aiyan et al. (2020)	6 例健康成人	言语想象 VS 言语产生(发音)	50 个 CVC (辅音−元音−辅音)单词，包含 /a/, /i/, /u/, /e/, /o/：can, car, cat, bad, dad, gas, lab, man, rat, tap; did, fit, kit, lip, pig, pin, rip, sim, sit, zip; bun, bus, cup, gum, hug, hut, jug, pip, sum, sun; bed, den, hen, her, jet, led, let, net, red, vex; box, cop, dog, fog, jog, lot, not, pot, rod, sob.	全脑区 (额叶、颞叶、顶叶、枕叶)	EEG	α 波、δ 波、θ 波、β 波、γ 波	①言语想象时，颞叶的 θ 波占主导；②言语产生(发音)时，额叶的 γ 波占主导
Nguyen et al. (2017)	15 例健康成人	言语想象状态下长单词、短单词和元音的对比	长单词：cooperate, independent 短单词：in, out, up 元音：/a/, /i/, /u/	全脑区	EEG	脑电波	脑活动信号确实均集中在位于布洛卡区上方的左前额叶，中额叶和顶叶，运动皮层和韦尼克区。想象时短单词和长单词在布洛卡区和韦尼克区出现时频差异；短单词比长单词更容易在高频段(31~70 Hz)和布洛卡区受到抑制。
郭苗苗等 (2018)	9 例健康成人	言语想象 VS 空闲期	4 个字：喝，右，吃，冷	全脑区	EEG	α 波、δ 波、θ 波、β 波、γ 波	想象“喝”时，F5 与 F6 电极的脑电信号能量在 9~16 Hz 频率区域，500 ms 到 2100 ms 时间范围内较基线有明显的增强。想象“右”时，EEG 信号在 1500 ms 之后，8~14 Hz 频率区域能量较基线明显增强。想象“吃”时，EEG 信号在 300~1800 ms 时间内，8~14 Hz 频率区域内能量较基线明显减弱。想象“冷”时，EEG 信号在 300~2000 ms 时间内，6~13 Hz 和 16~22 Hz 频率区域内能量较基线均有明显减弱。

续表

研究	被试	想象对比	内容	覆盖脑区	工具	信号	结果
Orpella et al. (2022)	21 例健康成人	言语想象 VS 阅读	3 个音节：/pa/，/ta/，/ka/	全脑区	MEG 面部 sEMG	脑磁信号	①起始阶段的视觉处理信息两种状态共享枕叶视觉皮层; ②言语想象时脑区激活的时间进程按照前 120 ms 的视觉区域活动，180 ms 时以颞外侧皮质(左侧)为主，在位置和时间上都与预期的语音编码一致。260~300 ms 时到达顶叶(听觉记忆皮层)、岛叶、双侧运动前区，反映听觉运动整合过程和发音运动计划过程。440 ms 后为广泛性左侧听觉皮层活动；阅读时以视觉信息解码为主; ③言语想象和阅读时均出现面部肌肉微弱电信号，不同音节/pa/、/ta/、/ka/之间的面部电流无差异。

3.3　想象句子时的脑神经信号

句子是人们日常交流最常使用的语言方式，对句子解码的研究具有更加深远的意义。目前关于句子想象的神经机制研究还较少，多是以想象时脑功能信号的识别率为主。Dash 等人(2020)尝试采用 MEG 探索了想象 5 个简单句子(例如：How are you? I need help.等)的脑功能信号的识别率，认为采用卷积神经网络模型(convolutional neural networks, CNN)脑信号解码识别率高达 93%。Lee 等人(2019)采用 EEG 识别想象句子的全脑区脑电信号的准确率为 34.2%。但两项研究均未汇报具体脑区的活动特征。进一步分析其研究的差异可能是由于脑功能神经信号采集技术和解码技术的差异引起的。MEG 对垂直于头皮表面(脑沟)的皮层信号最敏感，而 EEG 对平行表面(脑回)的信号敏感，且 MEG 可以很好的平衡时间和空间分辨率，这均有助于脑神经信号的采集和解码(Fyshe, 2020; Orpella et al., 2022)。在解码技术上，Lee 的研究采用公共空间特征提取方法和线性判别分析的分类方法，判断句子的十三分类范式识别的正确率(Lee et al., 2019)，而 Dash 的研究采用小波变换的特征提取和卷积神经网络计算五分类范式的识别率(Dash et al., 2020)，十三分类相对五分类更复杂。然而，由于 MEG 价格昂贵和不便携的特性，当前言语想象的主要技术手段还是以 ECoG 和 EEG 为主，但 MEG 研究的实验范式和解码方法值得借鉴。

综上所述，想象元音和音节时大脑的活动与实际言语产生相似，然而想象有意义的单词、句子时，则与实际言语产生有不同的脑电信号，语料的复杂性、词的长度、意义等会影响想象时的脑区活动。另外，信号较弱的脑区识别、刺激方式、脑信号检测技术等均会影响脑神经信号的识别。

4　总结与展望

综上所述，本文主要从两个方面梳理了言语想象的神经机制：一是言语想象与言语产生神经机制的异同；二是不同想象内容脑神经信号的特征。言语想象与实际言语产生在理论模型、激活脑区、神经传导路径等方面均有较多的相似性。然而，言语想象神经机制的研究还面临很多挑战。在 BCI 领域的研究中，对元音想象和音节想象的脑神经信号解码率较高，而对单词、辅音和句子方面的研究还处于初级阶段，想象具有语义性质的单词与实际言语产生时的大脑功能活动存在差异，且想象不同长度和语义的单词时大脑活动加工也存在差异。另外，相关神经损伤会一定程度上影响言语想象的能力，言语想象干预可以

改善患者的言语功能。虽然言语想象的神经机制研究已经取得了显著的成果，但在如下方面值得进一步探索。

4.1 探索评价言语想象质量的工具及神经解码范式

依据前述言语想象理论模型研究的不足，评价言语想象质量的工具及神经解码范式还需进一步探索。针对言语想象没有明确的行为表现的问题，后续可以通过想象质量问卷或想象内容转录等方法判断受试者想象的质量。各种形式的想象质量问卷已在运动想象的基础和行为研究中普遍应用(白学军 等, 2016; 刘华 等, 2017)，强调对动作想象的清晰度进行判断。言语想象同时包括“想象听”和“想象说”，以及会涉及到语义的内容，关于言语想象的质量问卷可能需要兼顾这些方面的内容。在神经损伤群体的研究中，需要提供可让其顺利进行言语想象的工具。在神经信号解码技术方面，已有研究多关注激活信号比较显著的脑功能区，而忽略了信号较弱的脑区的功能活动，这些区域的解码也许可以为解释言语想象加工机制做贡献。参考脑神经可塑性相关理论，还可增加相应脑损伤后可能的代偿区域的神经解码(Wilson et al., 2020)。BCI 的主要目的是改善神经损伤后患者的大脑与外界沟通和控制，关注患者言语想象时代偿脑区的神经解码有助于 BCI 的研究。例如通过改进脑神经信号解码的算法来解码失语症患者颞顶枕交界处、颞前腹区的神经信号(Gajardo-Vidal et al., 2021)，虽然在健康人群中该区域激活不明显，但可以有助于失语症患者 BCI 解码的正确率。

4.2 言语想象的神经传导路径可进一步扩展到脑控制回路、激活通路的研究

言语产生过程中神经传导路径包括控制回路(the control circuits)、直接激活通路(the direct activation pathway)、间接激活通路(the indirect activation pathway)和外周神经的最后共同通路(the final common pathway) (Duffy, 2013)，而当前关于言语想象的神经传导路径只涉及了额叶与颞叶之间的联系，以及面部肌电的检测。目前已有采用DTI追踪脑白质纤维束通路方面的研究证明，运动想象治疗后脑卒中患者患侧大脑白质纤维束的各向异性系数(fractional anisotropy, FA)升高，说明运动想象可以促进患者受损脑白质纤维的恢复(杨帆 等, 2017; Li et al., 2018)，后续可采用 DTI 技术结合合理的实验设计，进一步探索言语想象过程中脑控制回路、激活通路的机制研究。虽然当前BCI技术也仅是涉及了脑皮层的信号识别，言语想象的神经传导路径研究或许能够为 BCI 技术探索新的路径。另外，言语想象激活的信号强度较弱，但激活信号相对较弱的脑区也可能对 BCI 的识别产生影响，尤其在初级运动皮层是否被激活及其激活程度还存在一定的争议，这可能与不同研究所采用的脑功能检测技术不同有关，后续可能需要进行不同检测技术之间的对照研究和严格的实验设计控制无关变量。虽然 MEG 在识别言语想象脑神经信号方面的功能优于 EEG (Dash et al., 2020)，但 MEG 价格昂贵难以大规模使用。近年来流行的一种新型无创的近红外脑功能成像(fNIRS)，相比 fMRI、PET 和 EEG, fNIRS 对运动不敏感且具备较高的时间分辨率(安娟, 牟海荣, 2019; 白学军 等, 2016; 高晨阳 等, 2022)，可以考虑作为言语想象研究的检测技术。

4.3 探索言语障碍群体的言语想象机制

当前言语想象神经机制的研究成果多以健康成人为研究对象，不能说明言语障碍群体的言语想象机制。诸如前述不同类型失语症患者言语想象的脑区活动与言语产生存在差异，口吃患者言语想象与言语产生行为学上的差异，均提示神经损伤患者的言语想象机制与健康人群有区别，目前关于障碍群体言语想象的机制研究还较少。除外失语症和口吃，运动性言语障碍也是常见的神经损伤导致的言语障碍群体，但目前较少有研究关注这类群体。运动性言语障碍是神经损伤导致言语产生的计划、编程和执行功能失调(Duffy, 2013)，使言语产生的呼吸、发声、共鸣、构音、语音所需的肌肉控制失调导致言语的可懂度降低，这是脑卒中、帕金森病、脑外伤、脑瘫、肌萎缩侧索硬化症和重症肌无力等神经疾病的常见伴随症状。这些患者是应用 BCI 技术的主要群体之一，基于以往研究结论局灶性脑区受损的言语想象功能仍有所保留，后续可以进一步探索特定言语障碍群体下的言语想象机制，进而为言语想象的 BCI 研究提供理论指导。另外，后续可以进一步探索言语想象在以上言语障碍群体中的随机对照研究，以及对比治疗前后脑神经信号在脑区激活、神经传导路径等方面变化的差异，为言语障

碍的康复治疗新技术提供理论和实践支撑。

4.4 探索赋有语义性质的词语和句子的言语想象的脑神经信号特征提取技术

赋有语义性质的言语想象研究多集中在 BCI 交互准确率方面, 言语想象会受到想象内容的复杂程度和语义的影响(Nguyen et al., 2017), 词语和句子的想象脑机制还处于初级阶段, 这是受多方面因素影响的。首先, 信息技术领域对于脑神经信号特征提取的技术需进一步探索。通过 EEG、fNIRS 提取了言语想象的脑神经信号后, BCI 技术的核心是脑神经信号的特征提取、分类与解码等数据处理技术(刘艳鹏 等, 2022), 目前对元音想象中二分类、三分类的各类算法的准确率较高, 对于有语义性质的词汇或句子的算法较少, 多分类的准确性较低(陈霏, 潘昌杰, 2020), 这些信息技术的发展无疑能够促进复杂言语想象的研究。其次, 言语想象的刺激方式会影响复杂语义言语想象的效果。当前言语想象的刺激方式多以听觉刺激、视觉刺激为主, 复杂语义的加工机制较为复杂, 后续研究中可以进一步丰富触觉(Yoo et al., 2003)、嗅觉(Djordjevic et al., 2005)、味觉等其他刺激形式来促进复杂内容言语想象的脑信号提取。言语想象涉及的记忆加工理论, 要求受试者有正确想象内容的经验, 而诸如脑瘫等先天性的神经源性言语障碍群体来说, 其并未有正确构音的经验, 可能需要清晰、明确的动作或语音示范才能有助于其完成言语想象(Lust et al., 2016), 这可能需要先制作矢状面和冠状面的动态发音模型, 给患者呈现正确的发音动作示范, 便于其想象正确的发音动作。

5 结论

言语想象的神经机制对大脑预处理机制和 BCI 技术均起到了重要的作用。言语想象在理论模型、激活脑区、神经传导路径等均与正常言语产生有较多的相似之处, 尤其是简单元音和无意义音的想象机制促进了 BCI 技术的发展。复杂想象内容、神经损伤群体的言语想象与正常言语产生差异较大, 但目前该方向的研究多集中在 BCI 领域的研究, 在神经心理学方向的研究还较少。总体来讲言语想象的神经机制研究取得了一定的成果, 未来研究在言语想象质量评价工具及神经解码范式、脑控制回路、激活通路、词语和句子想象的脑神经信号、言语障碍群体的言语想象机制等方面进一步探索, 为有效提高 BCI 的识别率提供依据, 为言语障碍群体的沟通提供便利。

参考文献

安娟, 牟海荣. (2019). 基于EEG和fNIRS的多模态脑机接口运动想象参数研究. *科学技术创新, 28*, 101–102.

白学军, 张琪涵, 章鹏, 周菘, 刘颖, 宋星, 彭国慧. (2016). 基于 fNIRS 的运动执行与运动想象脑激活模式比较. *心理学报, 48*(5), 495–508.

蔡笑, 张清芳. (2020). 言语运动系统中前馈和反馈控制整合加工的作用机制. *心理科学进展, 28*(4), 588–603.

陈霏, 潘昌杰. (2020). 基于发音想象的脑机接口的研究综述. *信号处理, 36*(6), 816–830.

窦智, 肖凯龙, 陈颖, 邓泽华. (2021). 运动想象疗法联合语言训练对脑卒中后失语症患者语言功能的影响. *中外医学研究, 19*(6), 154–156.

高晨阳, 修嘉, 李婷. (2022). 运动想象任务下近红外光谱脑激活程度适应性研究. *中国激光, 49*(5), 204–212.

高晴, 陈华富. (2010). 基于动态因果模型的运动执行和运动想象脑网络研究. *电子科技大学学报, 39*(3), 457–460.

郭苗苗, 齐志光, 王磊, 徐桂芝. (2018). 语言脑机接口康复系统中的参数优化研究. *信号处理, 34*(8), 974–983.

刘华, 周军, 杨少峰, 史清钊, 程钰琦, 李洋, ... 张通. (2017). 中文版运动觉-视觉想象问卷在脑卒中患者中的信度. *中华物理医学与康复杂志, 39*(4), 259–263.

刘艳鹏, 龚安民, 丁鹏, 赵磊, 钱谦, 周建华, …伏云发. (2022). 基于言语想象的脑机交互关键技术. *生物医学工程学杂志, 39*(3), 596–611.

吴金香. (2019). 运动想象在脑卒中运动性失语患者中的应用研究. *护士进修杂志, 34*(6), 533–535.

吴拾瑶, 随力, 杨兰, 刘亮. (2021). 运动想象重塑脑功能的研究进展. *中国医学物理学杂志, 38*(11), 1449–1452.

杨帆, 桑德春, 张晓钰, 卢利萍, 褚宏宇. (2017). 运动想象疗法对脑卒中患者运动功能康复的效果. *中国康复理论与实践, 23*(9), 1081–1085.

杨晓芳, 江铭虎. (2014). 基于汉语音位发音想象的脑机接口研究. *中文信息学报, 28*(5), 13–23.

姚婧璠, 徐成, 陈红燕, 王铄, 张玉梅. (2021). 卒中后失语语言和非语言认知功能相关的静息态功能磁共振成像研究. *中国卒中杂志, 16*(3), 251–258.

张清芳, 钱宗愉, 朱雪冰. (2021). 汉语口语词汇产生中的多重音韵激活: 单词翻译任务的 ERP 研究. *心理学报, 53*(1), 1–14.

张清芳, 杨玉芳. (2003). 言语产生中的词汇通达理论. *心理科学进展, 11*(1), 6–11.

Ackermann, H., & Hertrich I. (2003). *Cerebellar contributions to speech motor control and auditory verbal imagery:*

Acoustic / kinematic analyses of ataxic dysarthria and functional magnetic resonance imaging in healthy subjects. In: Solé M. J., Recasens D., & Romero J. (Eds), *Proceedings of the 15th International Congress of Phonetic Sciences* (volume 1, pp. 163–167). Causal Productions, Adelaide.

Alderson-Day, B., & Fernyhough, C. (2015). Inner speech: Development, cognitive functions, phenomenology, and neurobiology. *Psychological Bulletin, 141*(5), 931–965.

André, A., Elia, F., Heidi, K., Peter, H., Haan, E. H. F. D., & Kahn, R. S. (2005). The functional neuroanatomy of metrical stress evaluation of perceived and imagined spoken words. *Cerebral Cortex, 15*(2), 221–228.

Arongna, Sakai, N., Yasu, K., & Mori, K. (2020). Disfluencies and strategies used by people who stutter during a working memory task. *Journal of Speech, Language, and Hearing Research, 63*(3), 688–701.

Barrows, P. D., Thomas, S. A., & Gordon, W. V. (2021). Assessing self-reported mood in aphasia following stroke: Challenges, innovations and future directions. *Journal of Stroke and Cerebrovascular Diseases, 30*(1), 105425–105425.

Birn, R. M., Bandettini, P. A., Cox, R. W., & Shaker, R. (1999). Event-related fMRI of tasks involving brief motion. *Human Brain Mapping, 7*(2), 106–114.

Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *NeuroImage, 32*(2), 821–841.

Bruno, V., Fossataro, C., & Garbarini, F. (2018). Inhibition or facilitation? Modulation of corticospinal excitability during motor imagery. *Neuropsychologia, 111*, 360–368.

Chengaiyan, S., Retnapandian, A. S., & Anandan, K. (2020). Identification of vowels in consonant–vowel–consonant words from speech imagery based EEG signals. *Cognitive Neurodynamics, 14*(1), 1–19.

Cooney, C., Folli, R., & Coyle, D. (2018). Neurolinguistics research advancing development of a direct-speech brain-computer interface. *iScience, 8*, 103–125.

Courson, M., & Tremblay, P. (2020). Neural correlates of manual action language: Comparative review, ALE meta-analysis and ROI meta-analysis. *Neuroscience & Biobehavioral Reviews, 116*, 221–238.

DaSalla, C. S., Kambara, H., Sato, M., & Koike, Y. (2009). Single-trial classification of vowel speech imagery using common spatial patterns. *Neural networks, 22*(9), 1334–1339.

Dash, D., Paul, F., & Wang, J. (2020). Decoding imagined and spoken phrases from non-invasive neural (MEG) signals. *Frontiers in Neuroscience, 14*, 290–304.

Djordjevic, J., Zatorre, R. J., Petrides, M., Boyle, J. A., & Jones-Gotman, M. (2005). Functional neuroimaging of odor imagery. *NeuroImage, 24*(3), 791–801.

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management.* St. Louis, MO: Elsevier.

Durand, E., Berroir, P., & Ansaldo, A. I. (2018). The neural and behavioral correlates of anomia recovery following personalized observation, execution, and mental imagery therapy: A proof of concept. *Neural Plasticity,* https://doi.org/10.1155/2018/5943759.

Ehrsson, H. H., Geyer, S., & Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *Journal of Neurophysiology, 90*(5), 3304–3316.

Fama, M. E., Hayward, W., Snider, S. F., Friedman, R. B., & Turkeltaub, P. E. (2017). Subjective experience of inner speech in aphasia: Preliminary behavioral relationships and neural correlates. *Brain and Language, 164*, 32–42.

Fama, M. E., Snider, S. F., Henderson, M. P., Hayward, W., Friedman, R. B., & Turkeltaub, P. E. (2019). The subjective experience of inner speech in aphasia is a meaningful reflection of lexical retrieval. *Journal of Speech, Language, and Hearing Research, 62*(1), 106–122.

Fyshe, A. (2020). Studying language in context using the temporal generalization method. *Philosophical Transactions of the Royal Society B, 375*(1791), 20180531.

Gajardo-Vidal, A., Lorca-Puls, D., Team, P., Warner, H., Pshdary, B., Crinion, J., … Price, C. (2021). Damage to Broca’s area does not contribute to long-term speech production outcome after stroke. *Brain, 144*(3), 817–832.

Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J. -C., & Warburton, E. A. (2011). The neural correlates of inner speech defined by voxel-based lesion-symptom mapping. *Brain, 134*(10), 3071–3082.

Gu, J., Zhang, H., Liu, B., Li, X., Wang, P., & Wang, B. (2019). An investigation of the neural association between auditory imagery and perception of complex sounds. *Brain Structure and Function, 224*(8), 2925–2937.

Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics, 72*(1), 43–53.

Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review, 102*(3), 594–621.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39*(5), 350–365.

Guenther, F. H., Brumberg, J. S., Wright, E. J., Nieto-Castanon, A., Tourville, J. A., Panko, M., ⋯ Kennedy, P. R. (2009). A wireless brain-machine interface for real-time

speech synthesis. *PLoS One, 4*(12), e8218.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96*(3), 280–301.

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics, 25*(5), 408–422.

Haire, C. M., Tremblay, L., Vuong, V., Patterson, K. K., Chen, J. L., Burdette, J. H., ⋯ Thaut, M. H. (2021). Therapeutic instrumental music training and motor imagery in post-stroke upper-extremity rehabilitation: A randomized-controlled pilot study. *Archives of Rehabilitation Research and Clinical Translation, 3*(4), 100162.

Heremans, E., Feys, P., Nieuwboer, A., Vercruysse, S., Vandenberghe, W., Sharma, N., & Helsen, W. (2011). Motor imagery ability in patients with early- and mid-stage Parkinson disease. *Neurorehabil Neural Repair, 25*(2), 168–177.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroence, 13*(2), 135–145.

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron, 69*(3), 407–422.

Houde, J. F., & Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in Human Neuroscience, 5*, 82–95.

Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the ecological validity of thinking on demand: Neural correlates of elicited vs. spontaneously occurring inner speech. *PLoS One, 11*(2), e0147932.

Jahangiri, A., & Sepulveda, F. (2017). The contribution of different frequency bands in class separability of covert speech tasks for BCIs. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBC), 2093–2096.

Jahangiri, A., & Sepulveda, F. (2018). The relative contribution of high-gamma linguistic processing stages of word production, and motor imagery of articulation in class separability of covert speech tasks in EEG data. *Journal of Medical Systems, 43*(2), 20–28.

Jones, S. R., & Fernyhough, C. (2007). Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness Cognition, 16*(2), 391–399.

Kearney, E., & Guenther, F. H. (2019). Articulating: The neural mechanisms of speech production. *Language, Cognition and Neuroscience, 34*(9), 1214–1229.

Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive processing: A canonical cortical computation. *Neuron, 100*(2), 424–435.

Kilteni, K., Andersson, B. J., Houborg, C., & Ehrsson, H. H. (2018). Motor imagery involves predicting the sensory consequences of the imagined movement. *Nature Communnications, 9*(1), 1617–1625.

Klingbeil, J., Wawrzyniak, M., Stockert, A., & Saur, D. (2019). Resting-state functional connectivity: An emerging method for the study of language networks in post-stroke aphasia. *Brain and Cognition, 131*, 22–33.

Langland-Hassan, P., Faries, F. R., Richardson, M. J., & Dietz, A. (2015). Inner speech deficits in people with aphasia. *Frontiers in Psychology, 6*, 528.

Laures-Gore, J., Stache, M., Moore, E., & Tullis, C. (2021). App-based data collection, mental imagery, and naming performance in adults with aphasia. *Complementary Therapies in Clinical Practice, 44*, 101422.

Lebon, F., Rouffet, D., Collet, C., & Guillot, A. (2008). Modulation of EMG power spectrum frequency during motor imagery. *Neuroscience Letters, 435*(3), 181–185.

Lee, S. H., Lee, M., Jeong, J. H., & Lee, S. W. (2019). *Towards an EEG-based intuitive BCI communication system using imagined speech and visual imagery.* Paper presented at the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy.

Li, F., Zhang, T., Li, B. J., Zhang, W., Zhao, J., & Song, L. P. (2018). Motor imagery training induces changes in brain neural networks in stroke patients. *Neural Regeneration Research, 13*(10), 1771–1781.

Lust, J. M., Wilson, P. H., & Steenbergen, B. (2016). Motor imagery difficulties in children with cerebral palsy: A specific or general deficit? *Research in Developmental Disabilities, 57*, 102–111.

Ma, S. Y., Jin, D. T., Zhang, M., Zhang, B. X., Wang, Y., Li, G., & Yang, M. (2019). *Silent speech recognition based on surface electromyography.* 2019 Chinese Automation Congress.

MacKay, D. G. (1992). *Auditory imagery*. New York: Psychology Press. https://doi.org/10.4324/9781315807621.

Martin, S., Iturrate, I., Millán, J. D. R., Knight, R. T., & Pasley, B. N. (2018). Decoding inner speech using electrocorticography: Progress and challenges toward a speech prosthesis. *Frontiers in Neuroscience, 12*, 422–422.

Naito, E., Kochiyama, T., Kitada, R., Nakamura, S., Matsumura, M., Yonekura, Y., & Sadato, N. (2002). Internally simulated movement sensations during motor imagery activate cortical motor areas and the cerebellum. *Journal of Neuroscience, 22*(9), 3683–3691.

Neef, N. E., Anwander, A., Bütfering, C., Schmidt-Samoa, C., Friederici, A. D., Paulus, W., & Sommer, M. (2018). Structural connectivity of right frontal hyperactive areas scales with stuttering severity. *Brain, 141*(1), 191–204.

Neef, N. E., Bütfering, C., Anwander, A., Friederici, A. D., Paulus, W., & Sommer, M. (2016). Left posterior-dorsal area 44 couples with parietal areas to promote speech fluency, while right area 44 activity promotes the stopping of motor responses. *NeuroImage, 142*, 628–644.

Netsell, R., & Bakker, K. (2017). *Fluent and dysfluent inner speech of persons who stutter: Self-report.* Unpublished manuscript. Missouri State University, Springfield, MO.

Nguyen, C. H., Karavas, G. K., & Artemiadis, P. (2017). Inferring imagined speech using EEG signals: A new approach using Riemannian manifold features. *Journal of Neural Engineering, 15*(1), 016002.

Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: The flexible abstractness of inner speech. *Memory & Cognition, 38*(8), 1147–1160.

Orpella, J., Mantegna, F., Assaneo, M. F., & Poeppel, D. (2022). Speech imagery decoding as a window to speech planning and production. *BioRxiv*, 494046.

Orpella, J., Mas-Herrero, E., Ripollés, P., Marco-Pallarés, J., & de Diego-Balaguer, R. (2021). Statistical learning as reinforcement learning phenomena. *PLoS Biology, 19*(9), e3001119.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M., & Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *NeuroImage, 14*(1), 182–193.

Parrell, B., Agnew, Z., Nagarajan, S., Houde, J., & Ivry, R. B. (2017). Impaired feedforward control and enhanced feedback control of speech in patients with cerebellar degeneration. *Journal of Neuroscience, 37*(38), 9249–9258.

Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience, 20*(10), 624–634.

Pei, X. M., Leuthardt, E. C., Gaona, C. M., Brunner, P., Wolpaw, J. R., & Schalk, G. (2011). Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *NeuroImage, 54*(4), 2960–2972.

Perrone-Bertolotti, M., Rapin, L., Lachaux, J. P., Baciu, M., & Lœvenbruck, H. (2014). What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. *Behavioural Brain Research, 261*, 220–239.

Proix, T., Delgado S, J., Christen, A., Martin, S., Pasley, B., Knight, R., ⋯ (2021). *Imagined speech can be decoded from low- and cross-frequency features in perceptual space*. Preprint version. DOI: 10.1101/2021.01.26.428315

Ramirez-Quintana, J. A., Macias-Macias, J. M., Ramirez-Alonso, G., Chacon-Murguia, M. I., & Corral-Martinez, L. F. (2021). *A novel deep capsule neural network for vowel imagery patterns from EEG signals.* Available at SSRN: http://dx.doi.org/10.2139/ssrn.3991518

Riaz, A., Akhtar, S., Iftikhar, S., Khan, A., & Salman, A. (2015). *Inter comparison of classification techniques for vowel speech imagery using EEG sensors.* The 2014 2nd International Conference on Systems and Informatics (ICSAI 2014), Shanghai, China.

Sandhya, C., Srinidhi, G., Vaishali, R., Visali, M., & Kavitha, A. (2015). *Analysis of speech imagery using brain connectivity estimators.* In 2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC). http://dx.doi.org/10.1109/ICCI-CC.2015.7259410.

Schumacher, R., Halai, A. D., & Lambon Ralph, M. A. (2019). Assessing and mapping language, attention and executive multidimensional deficits in stroke aphasia. *Brain, 142*(10), 3202–3216.

Scott, M., Yeung, H. H., Gick, B., & Werker, J. F. (2013). Inner speech captures the perception of external speech. *The Journal of the Acoustical Society of America, 133*(4), EL286–EL292.

Sereshkeh, A. R., Yousefi, R., Wong, A. T., & Chau, T. (2018). Online classification of imagined speech using functional near-infrared spectroscopy signals. *Journal of Neural Engineering, 16*(1), 016005.

Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychological Medicine, 31*(2), 241–253.

Solms, M. (2000). Dreaming and REM sleep are controlled by different brain mechanisms. *Behavioral and Brain Sciences, 23*(6), 843–850.

Souto, D. O., Cruz, T., Fontes, P., & Haase, V. G. (2020). Motor imagery in children with unilateral cerebral palsy: A case–control study. *Developmental Medicine & Child Neurology, 62*(12), 1396–1405.

Stark, B. C., Geva, S., & Warburton, E. A. (2017). Inner speech's relationship with overt speech in poststroke aphasia. *Journal of Speech, Language, and Hearing Research, 60*(9), 2406–2415.

Tak, S., Kempny, A. M., Friston, K. J., Leff, A. P., & Penny, W. D. (2015). Dynamic causal modelling for functional near-infrared spectroscopy. *NeuroImage, 111*, 338–349.

Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology, 1*, 166–188.

Tian, X., & Poeppel, D. (2012). Mental imagery of speech: Linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience, 6*, 314–323.

Tian, X., & Poeppel, D. (2015). Dynamics of self-monitoring and error detection in speech production: Evidence from mental imagery and MEG. *Journal of Cognitive Neuroscience, 27*(2), 352–364.

Tian, X., Zarate, J. M., & Poeppel, D. (2016). Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex, 77*(7), 1–12.

Torres-García, A. A., Reyes-García, C. A., Villaseñor-Pineda, L., & García-Aguilar, G. (2016). Implementing a fuzzy inference system in a multi-objective EEG channel selection model for imagined speech classification. *Expert Systems with Applications, 59*, 1–12.

Wang, L., Zhang, X., Zhong, X., & Zhang, Y. (2013). Analysis and classification of speech imagery EEG for BCI. *Biomedical Signal Processing and Control, 8*(6), 901–908.

Weber, R. J., & Bach, M. (1969). Visual and speech imagery. *British Journal of Psychology, 60*(2), 199–202.

Wilson, G., Stavisky, S., Willett, F., Avansino, D., Kelemen, J., Hochberg, L., ··· Shenoy, K. V. (2020). Decoding spoken English from intracortical electrode arrays in dorsal precentral gyrus. *Journal of neural engineering, 17*(6), 066007.

Yoo, S. S., Freeman, D. K., McCarthy, J. J., & Jolesz, F. A. (2003). Neural substrates of tactile imagery: A functional MRI study. *Neuroreport, 14*(4), 581–585.

Zhang, X., Li, H., & Chen, F. (2020). EEG-based classification of imaginary Mandarin tones. *Annual International Conference of the IEEE Engineering in Medicine & Biology Society* , 3889–3892.

Neural mechanism of speech imagery

WANG Yongli[1], GE Shengnan[2], Lancy Lantin Huang[3], WAN Qin[1], LU Haidan[1]

([1] *Department of Rehabilitation Sciences, Faculty of Education, East China Normal University, Shanghai* 200062, *China*)

([2] *School of Life Sciences, East China Normal University, Shanghai* 200241, *China*)

([3] *Bright Speech International Trade Co., Ltd. Shanghai* 200131, *China*)

Abstract: Speech imagery not only plays an important role in the pre-processing mechanism of the brain, but it is also the latest technology in the field of brain-computer interface (BCI) research. The theoretical model of speech imagery, activation of brain regions, and neural conduction pathways all have many similarities with speech production. However, the brain activation mechanism of speech disorders, as well as imagining contents with complex semantics is different from speech production. Given the complexity of the human speech system, research on the neural mechanisms of speech imagery still faces a series of challenges. Future research can be focused on evaluation tools of speech imagery quality and neural decoding paradigm, brain control circuits, activation pathways, speech imagery mechanism of speech impaired groups, and brain neural signals of word and sentence imagination. Further exploration will help provide a basis for effectively improving the recognition rate of BCI and facilitating communication in speech disorders.

Keywords: speech imagery, speech motor planning, brain-computer interface, neural mechanism